电子断层三维重构技术的研究与进展

张法，万晓华，储琪

摘要：电子断层三维重构技术能够在纳米尺度下重构出不具有全同性的细胞或大分子的三维结构，已经成为一种公认的研究生物大分子结构强有力手段。然而，电子断层三维重构仍然存在二维图像对位不精确、重构算法精度低和重构速度极其缓慢等问题。针对以上问题，本文详细介绍了电子断层三维重构的发展和现状，分析了当前所面临的主要挑战问题，着重介绍了我们在电子断层三维重构相关研究工作上的进展。

关键词：电子断层 三维重构 对位算法 迭代算法 并行计算 GPU

1 引言

结构生物学是研究生物大分子特定三维结构及结构的特定运动、结构相关生物学功能的一门学科，其起源可追溯到上世纪 50 年代沃森（J. Waston）和克里克（F. Crick）发现 DNA 双螺旋结构，60 年代佩鲁茨(M. Perutz)利用 X 射线晶体衍射技术获得肌球蛋白的三维结构[1]。现在结构生物学在分子生物学研究中占据了主流地位，已成为药物设计、疫苗开发等研究的基础。

众所周知，蛋白质是生命活动的体现者之一，对蛋白质三维结构的研究一直是结构生物学的主要研究内容。目前 X 射线晶体衍射技术、核磁共振技术和电镜三维重构技术是结构生物学的三大研究手段。其中 X 射线晶体衍射技术是目前分辨率最高的结构测定方法，可以获得生物分子的原子分辨率三维结构，但是生物大分子的晶体难以获得，结构解析极其困难；核磁共振技术可以获得蛋白质在溶液中的三维结构，能够获得蛋白质的动态变化信息，但是其研究对象的分子量通常难以超过 20kD，因此这两种方法，极大限制了人类对蛋白质功能以及生命活动规律的认识。随着生物样品制备技术的不断完善，电子显微镜设备的进步以及数字图像处理技术的发展，电镜三维重构技术，特别是冷冻电镜三维重构技术，能够解析那些无法应用 X 射线晶体衍射和核磁共振技术进行分析的蛋白质，弥补这两种技术的不足，已成为研究生物大分子三维结构的重要手段。萨利（A. Sali）等在《Nature》提出了今后结构生物学研究的发展方向，即将包括 X 射线晶体衍射、核磁共振、冷冻电镜三维重构、荧光能量共振等多种生物物理技术联合起来，在分子水平上理解细胞的结构，其中冷冻电镜三维重构技术起到了重要的纽带作用[2]。

电镜三维重构技术的思想最早由德罗西耶（D. DeRosier）和克卢格（A. Klug）在 1968 年提出。上世纪 80 年代，快速冷冻生物样品的制备技术和低温电子显微镜技术的快速发展与实用化，为利用电镜三维重构技术研究蛋白质三维结构奠定了基础。当前，利用电镜三维重构技术解析蛋白质超分子复合物的结构已分化发展为三种具有不同特点和适用范围的方法：电子晶体学（Electron Crystallography）、单颗粒三维重构（Single Particle）和电子断层三维重构（Electron Tomography，ET）。其中，电子晶体学主要用于解析蛋白质分子的二维晶体样品；单颗粒三维重构，主要解析具有全同性的蛋白质分子的三维结构，最新研究成果已经可以在原子分辨率下解析具有高对称性的病毒分子三维结构[3]（关于单颗粒三维重构的详细介绍，参见本刊第 56 期“单颗粒冷冻电镜三维重构研究的新进展”一文[4]）；电子断层三维重构通过获取同一区域多个角度的投影图来反向重构所研究对象的三维结构，适合于在

纳米级尺度上研究不具有结构均一性的蛋白、病毒、细胞器以及它们之间组成的复合体的三维结构。由于无需分子样品具有结构同一性和对称性，因此，虽然目前电子断层三维重构所获得结构的分辨率（约 4~10 纳米）不能与其它两种电镜三维重构方法相比，但其在研究非定形、不对称和不具有全同性生物样品的三维结构和功能中有着不可替代的重要作用。此外，电子断层三维重构的适用尺度非常广泛，包括从分子水平的蛋白质，到亚细胞水平的细胞器，以至细胞水平的组织结构，可以有效弥补 X 射线晶体衍射、核磁共振、单颗粒三维重构等方法得到的高精度结构与光学显微技术得到的低分辨率细胞整体图像之间的空白，为深入理解生物大分子的功能提供重要而有益的信息。因此，电子断层三维重构技术被《科学》杂志评为 2002 年十大科技突破之一[5]，近年来在冷冻电镜研究领域得到了越来越多的关注。

2 电子断层三维重构技术

电子断层三维重构技术（ET）是从一个物体的电子显微投影图像重构物体内部结构的技术。类似于人们熟知的 CT 扫描技术，通过获取同一个物体的多个连续角度下的二维投影图像来反向重构它的三维结构。其主要策略是将制备好的生物样品放在透射电子显微镜中，让样品沿着一个与电子束垂直的轴旋转，每旋转一定角度，拍摄一张样品在相对应方向上的二维投影图像，然后对这一系列二维投影图像进行对位，利用三维重构算法得到样品的三维结构。和 CT 扫描技术转动光源获得不同角度的投影图不同，ET 的光源（电子束）固定，通过倾转生物样品来得到不同方向的投影。根据电镜设备的不同，样品倾转范围大概在±60°~±80°之间，间隔角度为 1°或 2°。因此，ET 数据存在信息缺失，即缺失楔问题（missing wedge）。电子断层三维重构过程主要包括以下几个步骤[6]：(1). 样品制备、(2). 获取二维投影图像、(3). 投影图像的对位、(4). 断层图像的三维重构、(5). 重构结果的后期处理，如图 1 所示[1]。

样品制备：电子断层重构技术主要研究细胞环境下的三维结构，与传统的CT扫描使用的X射线相比，电子束能穿透的样品厚度较小，限制了电子断层技术对大的生物样品（如完整的真核细胞）的研究。对于电子束难以穿透的厚样品，如细胞等，需要采用切片的方法进行制样，经过脱水、染色后，进行观察和拍摄。对于电子束能穿透的较小或较薄的样品，如病毒、细菌等，可以采用快速冷冻的方法准备样品并进行冷冻电子断层三维重构。冷冻电子断层三维重构的主要优势是样品接近于天然状态，克服了由于脱水和染色等可能带来的假象，较好地保护了生物样品结构的完整性。

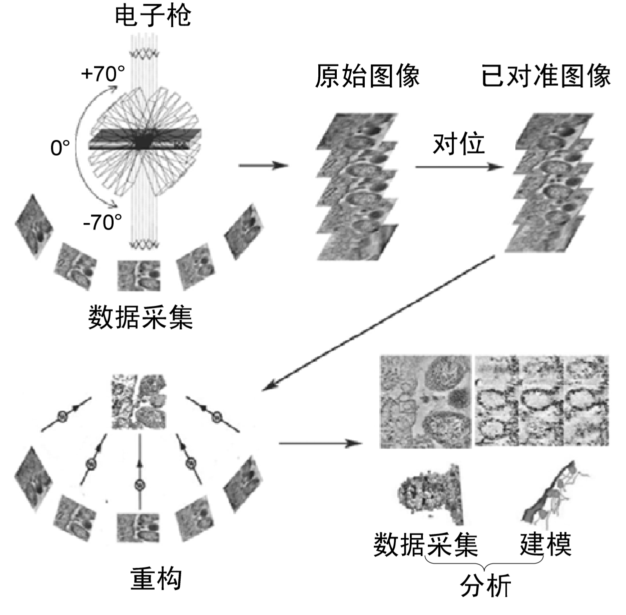


图 1. 电子断层三维重构流程示意图。

二维投影图像的获取：电子断层的二维投影图像收集首先是对样品选取合适的成像区域，然后对该区域连续收集不同角度的成像投影照片。电子断层二维投影图像的收集过程大概如下：(1). 寻找合适区域，为了减少对样品成像区域的电子辐射损伤，一般利用低电子束剂量模式的数据采集方案；(2). 调节样品

台旋转中心高度，使样品台旋转时样品维持在成像中心范围；(3). 调焦；(4). 找到同一成像区域的精确跟踪；(5). 照相；(6). 样品倾斜到下一个角度，重新找到样品的成像区域中心，然后重复以上四步：寻找、调焦、跟踪和照相。

由于电子断层在收集二维投影图像时，需要对样品同一区域反复成像，往往容易造成较大幅度的辐照损伤，因此必须严格控制样品台的精确倾转，对成像区域的自动跟踪，以及每次成像的电子束剂量等因素。在获取投影图像的过程中，样品在旋转时由于机械原因，往往会有一定的偏移，需要对不同倾斜角度的图像通过互相关运算来完成精确跟踪。为了提高图像衬度和跟踪的准确性，可以在样品中加入胶体金颗粒。目前已有一些自动化的电子断层数据收集软件，比如 UCSF tomography[7]，FEI 公司的 Explor3D 等。这些软件能够利用最初几张图像的偏移算出整个数据收集过程中图像的偏移轨迹，从而大大降低了额外增加的剂量对样品的损伤，并且能够比较精确地控制电镜和 CCD 相机，完成自动聚集功能。

二维投影图像的对位：由于在收集二维投影图像时，样品在倾斜过程中不可避免地会出现倾转轴改变、样品本身移动、投影本身移动、投影拍摄时CCD照相机参数不一致等问题，使得投影的系列二维图像之间发生偏移、旋转或者放大缩小等效应。因此，在三维重构之前，必须对一个倾斜系列投影的图像进行对位，首先是计算倾斜过程中由于样品台等的移动造成的图像中心的移动；其次，还需要对由于现存的成像条件导致的倾斜角度，图像本身的旋转、扭曲以及放大倍数等进行修正。二维投影图像对位的质量直接决定了最终重构结果的质量。

目前，二维投影图像对位常用的方法主要有基于标记点的方法和基于图像互相关匹配的方法。基于标记点的方法是在样品制备过程中，加入若干一定大小的胶体金（golden particles）作为标记物，然后根据胶体金标记的对应，建立投影模型，求解偏移、旋转或放大缩小等对位参数。这种方法常用的软件是 IMOD[8]。但这种方法所添加的胶体金极有可能造成样品污染，而且在倾转过程中胶体金可能会发生漂移，导致无法很好地对位。另一种对位方法是利用图像本身的互相关分析对图像进行对位，其中比较常用的软件有 Protomo[9]。除了 IMOD 和 Protomo 外，还有其他一些软件用于二维图像的对位，如 Spider[10]等以及 FEI 的 Inspec3D 等商业软件，但是精确的二维图像对位一直是该领域的一个难点问题。

电子断层三维重构：电子断层的三维重构是由所有对位的二维投影图像重新反向投影到一个3D体积中形成的，其算法主要有两种：背投影重构和迭代重构。

背投影重构算法包括直接背投影（back projection，BP）、滤波背投影（filter back projection, FBP）和加权背投影（weighted back projection，WBP）。直接背投影算法的基本思想是三维物体某一点的密度等于通过该点所有射线密度的总和。直接背投影算法是一种最简单的重构算法，但它容易使重构结果在实空间产生点扩散的不良效果。滤波背投影算法提出对重构结果进行点扩散函数的去卷积处理，但由于样品的空间几何参数未知，无法得到有效的点扩散函数。加权背投影算法则是对二维投影图像傅里叶变换的不同频率进行加权处理，然后在实空间中进行反投影操作。迭代重构算法完全是在实空间进行，建立三维物体中所有点的密度与投影图中所有点的密度之间的线性方程组，通过迭代方法求解线性方程组得到三维模型，使得原始投影图和重构投影图尽量接近。常用的迭代重构算法有代数重构算法[6]（Algebraic Reconstruction Technique，ART）、同步代数重构算法[11]（Simultaneous Algebraic Reconstruction Technique，SART）和联合迭代重构算法[12]（Simultaneous Iterative reconstruction Technique，SIRT）。迭代重构算法具有一致的权函数，但计算量巨大，为了得到满意的重构结果往往需要花费大量时间。

重构结果的后期处理：为了减少对样品的损伤，对样品拍照时往往电子束剂量很少，所

以得到的二维投影图像的信噪比都比较低，由此产生的重构结构噪音也很大。所以对重建结果要进行适当的降噪处理以突出研究对象。常用的去噪方法包括中值滤波、非线性滤波等。其中，各向异性扩散的非线性滤波方法效果较好。为了进一步将研究对象从背景和噪声中区分开，需要对重建结果进行分割。最常用的方法是使用手动分割，需要用户绘制对象和背景的边界轮廓。分割完成后可以利用表面渲染技术对重构结果进行可视化。常用的分割软件包括IMOD[8]和VAT4M[13]。

2.1 当前研究面临的问题

近年来电子断层技术取得了长足的进展，已经成为结构生物学研究中至关重要的研究手段。我国在电子断层重构领域的研究也取得了迅猛的进展。国内已拥有了包括 2 台 Titan Krios、1 台 Tecnai F30 等在内的高端电子显微镜研究设备。同时，多家研究机构也计划或正在购置更多的电镜设备。国内已经具备了很好的契机来开展电子断层成像方面的工作。然而与迅猛发展的硬件设施相比，我们在相应软件方面的研究就显得严重滞后，存在以下三方面的问题，制约着电子断层重构技术的进一步发展：

1. 缺乏准确的二维投影图像对位算法

如前所述，在电子断层成像对样品投影成像过程中，各种因素会导致最终生成的图像序列之间出现偏移、旋转或者放大缩小等效应，所以在三维重构前需要对系列二维投影图像进行对位。

目前，最常用的对位方法是前述之胶体金标记法和无标记对位方法。胶体金标记法存在很多缺点，所以需要更普适的无标记方法来对位投影图像。无标记对位方法主要有以下几类：(1). 互相关方法，其缺陷在于不同角度下的投影成像有可能差别很大，计算的偏移值不准确，且容易造成累积误差；(2). 等价线方法，其局限性在于只能求解一个对位参数，无法准确地对位投影图像；(3). 迭代对位方法，需要进行一次三维重构和求解参数进行对位操作，计算量巨大；(4). 基于特征匹配的对位方法，使用互相关方法来匹配特征点，存在着由于互相关方法所带来的点与点之间对应不精确的问题。因此，电子断层技术迫切需要更为精确的二维投影图像对位方法。

2. 缺乏高精度的三维重构算法

电子断层三维重构算法主要包括加权背投影和迭代算法两种。加权背投影是由雷德麦彻（M. Rademacher）提出的最早的电子断层重构算法。它利用中央截面定理，其基本原理是将每个二维投影图像背向其倾斜角方向投影到三维空间中，多个背投影图像在三维空间中叠加形成样品的三维结构。在重构过程中，通过权重滤波器实现取样平均的目的。由于加权背投影算法简单便于实现，目前流行的大部分主流电子断层重构软件，如 IMOD[8]，SPIDER[10] 等均采用加权背投影算法进行三维重构。然而，加权背投影算法容易受到信息缺失的影响，导致 z 轴方向上的重构结果错误，此外加权背投影算法在重构过程中没有考虑噪声带来的影响，使得重构结果不精确。

近几年人们开始考虑采用迭代算法进行电子断层三维重构。由于电子断层中投影方向只能覆盖有限角范围，并且投影图片噪声很大，迭代算法在这种数据不足和高噪声的条件下能够获得较好的重构结果。然而，传统的迭代算法具有以下缺陷：(1). 往往任意选取重构像素值的初值，通常情况下选取全 0（或全 1）向量，可能偏离真实值较远，因此需要多次迭代才能达到所需的收敛值； (2). 在更新投影数据时采用的访问方式是顺序访问策略，由于相邻投影方向之间有很大的相关性，容易引入误差，影响重构结果的精度，并且降低迭代算法

的收敛速度；(3). 迭代算法在每次迭代更新时，误差的补偿只考虑加权矩阵，这显然是不合理的，因为投影误差是由加权矩阵和当前像素值共同决定的。

3. 缺乏高效的三维重构并行算法

随着电子显微技术的飞速发展，为了能够得到更高分辨率的重构图像，往往需要增加投影图像的数量和尺寸，这样会导致电子断层三维重构需要大量的计算资源和处理时间。然而，面对这些大数据量的图像，现有的方法无法进行及时有效的处理，已经成为限制电子断层三维重构广泛应用的一个瓶颈问题。围绕高性能计算技术降低三维重构时间，人们进行了大量的研究[14-16]。当前大部分电子断层三维重构软件，如 IMOD[8]，SPIDER[10]，都实现了传统并行计算架构（如集群环境）上的算法。然而传统集群系统的昂贵造价以及 GPU 硬件技术的迅猛发展使得利用 GPU 开发高效的电子断层三维重构算法成为当前研究的热点[17,18]。

3 我们的工作

针对以上述目前电子断层存在的问题，本节主要介绍我们课题组在电子断层重构方面的一些研究工作。首先，介绍在电子断层中一种新的基于特征提取的二维投影图像对位算法 SAMA；其次，介绍一种收敛速度更快，重构结果更精确的迭代重构算法 ASART；最后，介绍了我们开发的国内首款基于 GPU 平台的电子断层重构软件 ATOM。

3.1 二维投影图像对位算法 SAMA

针对当前无胶体金样品的二维投影图像对位问题，我们提出了一种基于特征匹配的无标记对位新方法—SAMA(SIFT-based Automatic Markerless Alignment)。该方法建立了样品在投影时三维密度点与投影点的关系的非线性模型，于是对位问题转化为根据测量样本估计模型参数的参数估计问题。通过在每幅二维投影图像中提取特征点作为测量样本点，根据测量样本的精确性来选择合适的非线性模型，并进行参数估计。此方法的框架包括：特征点提取、特征点匹配、特征点跟踪以及模型选择与参数求解 4 个部分。SAMA 方法的具体步骤如下[19]：

特征点提取：图像的特征点通常包括两个部分：(1). **表示子**（descriptor） 即特征点独特的位置，一般根据图像数学特征进行提取，然后以坐标方式显式表达；(2). **描述子**（description） 表示子领域内的信息，用于匹配不同的表示子。早前的特征提取方法主要关注于如何设计具有区分度的表示子，较少涉及如何设计具有区分度的描述子。我们的特征点提取采用简化的尺度不变特征变换（Scale Invariant Feature Transform，SIFT)算法，对表述子和描述子都进行加强。其中表示子是同一图像不同尺度中的梯度稳定点；描述子是一个 128 维向量来描述表示子周围灰度的统计信息和灰度方向信息。SIFT 算法是目前最为精确的图像特征点提取方法，简化的 SIFT 方法能显著提高其速度，但不影响其精确性。

特征点匹配：由于图像上的每个特征点的信息都可以用一个 128 维向量描述子表示，且不同特征点的描述子区别很大，两幅图像之间可通过计算最小欧式距离的方法进行特征点匹配。此外两幅图像可以根据相隔角度估计出一幅图像上的特征点在另一幅图像上的位置差别的最大值，在此最大位置差别区域之内搜索与之欧式距离的最小点作为匹配特征点。若此最小欧式距离小于一定阈值，则将此匹配记录下来。最后对所有图像两两之间计算特征点匹配，并利用两幅投射图像之间存在着二阶几何张量关系，去除错误匹配对。

特征点跟踪：通过对图像中正确的匹配对进行串联来形成特征链。每条特征链表示同一三维密度点的投影序列。初始时，对一幅图像中的特征点 $\boldsymbol{a}$ 找到其在其它图像中的匹配点 $\boldsymbol{b}$，在其它图像中如果 $\boldsymbol{b}$ 又与另外图像有匹配点 $\boldsymbol{c}$，则将此串联起来。这样对每个图像对，都搜

索原有的特征链，若匹配对对应在原有某条特征链上，则此链上加点，否则重新生成新特征链。

模型选择与参数估计: 求解电子断层三维重构中三维密度点与二维投影点的关系模型可表示为：$\hat{\boldsymbol{m}}_{ij}=\boldsymbol{A}_j\boldsymbol{r}_i+\boldsymbol{t}_j=s_j\boldsymbol{R}_z(\alpha_j)\boldsymbol{P}\boldsymbol{R}_y(\vartheta_j)\boldsymbol{R}_x(\varphi_j)\boldsymbol{r}_i+\boldsymbol{t}_j$。我们需要估计 $s_j,\alpha_j,\vartheta_j,\varphi_j,r_i,t_j$ 这 6 个参数。其中 $\hat{\boldsymbol{m}}_{ij}$ 表示第 j 幅图像中对应的二维投影点坐标， $\boldsymbol{A}_j$ 表示从三维坐标到二维坐标的投影矩阵，$\boldsymbol{r}_i$ 表示第 i 个三维密度点，t_j 为成像时二维投影点的平移向量，$\alpha_j,\vartheta_j,\varphi_j$ 分别为绕 z，y，x 轴的倾转角度。

在参数估计过程中我们使用捆绑调整(bundle adjustment)和鲁棒的异常点（outlier）去除方法，根据误差特性对参数进行估计。在得到上述 6 个参数后，对每幅图像进行对位。

使用 SAMA 方法对真实电子断层成像数据进行对位，参数估计残差值在 0.2~0.4 像素之间，与使用胶体金标记的对位方法进行对位的残差值结果（0.2~0.7 像素）具有可比性，且经过此方法对位后重构的图像与使用胶体金方法对位重构的图像效果相当。

3.2 三维迭代重构算法 ASART

迭代算法认为重构的三维结果 $\boldsymbol{f}$ 可以看成一组基函数 $\boldsymbol{b}$ 和像素灰度值 $\boldsymbol{x}$ 的线性组合，表示如下：

$$f(r)\approx\sum_{j=1}^{J}x_j b_j(r)$$

其中，$\boldsymbol{J}$ 为三维空间像素点的个数。传统的体素模型法（voxel model）以单个小立方体，即体素（voxel），作为三维图像的基本单元，使用的基函数为分段函数，无法正确地体现重构样品密度空间的连续性。我们采用一种 blob 模型，认为整个三维空间由多个全同的球体构成（如图 2 所示），由于球体之间相互交叉重叠，能够更好地体现密度空间的连续性，使重构结果更加精确。

根据电子断层的投影原理，每条射线的投影值可以看成该射线经过的像素点灰度值的加权和，用如下线性系统表示：

$$p_i=\sum_{j=1}^{J}w_{ij}x_j,$$

其中，$\boldsymbol{p}_i$ 表示第 $\boldsymbol{i}$ 条射线的投影值，$\boldsymbol{w}_{ij}$ 表示加权因子，反映了像素 $\boldsymbol{x}_j$ 对射线投影 $\boldsymbol{p}_i$ 的贡献。在这一线性模型下，电子断层重构的目的是采用迭代算法根据测量所得的 $\boldsymbol{p}_i$ 值和可以计算得到的 $\boldsymbol{w}_{ij}$ 值估计各个像素点灰度值 $\boldsymbol{x}_j$，最后得到整个三维结构 $\boldsymbol{f}$。在 blob 模型下，由于每个球体的投影值跟射线方向无关，我们利用一种印迹（footprint）方法求解 $\boldsymbol{w}_{ij}$，先提前计算单个球体在不同 $\boldsymbol{r}$ 位置上的投影值并存在列表中，然后根据像素点 $\boldsymbol{x}_j$ 到射线 $\boldsymbol{p}_i$ 的距离查表得到对应的 $\boldsymbol{w}_{ij}$ 值。

图 2. Blob 重构模型

在求解 $\boldsymbol{X}$ 值时利用现有的迭代算法仍然存在收敛速度较慢，结果不精确等不足，因此我们提出一种自适应联合代数迭代算法 ASART[20]，采用以下四种关键技术来提高迭代算法的收敛速度，获得高精度的重构结果。

- **背投影算法（BPT）**

在现有迭代算法中，大部分 $X^{(0)}$的选取是任意的，通常情况下选取全 0（或全 1）向量，因此要使所求像素值 X 达到收敛可能需要迭代多次。为了能够减少迭代次数，加快算法收敛速度，我们考虑采用直接背投影算法（BPT）估计像素点初值 $X^{(0)}$。直接背投影算法是一种简单的重构算法，它认为每个像素点的灰度值是所有通过该点的射线投影加权之和。因此，像素点灰度初值 $X^{(0)}$可表示为：

$$x_j^{(0)}=\sum_{i=1}^{M}\frac{w_{ij}p_i}{\sum_{i=1}^{M}w_{ij}},(j=1,2,\ldots,N)\ 。$$

虽然背投影算法估计像素灰度初值会引入星状伪迹，但比任意选取初值的方法要更接近真实的灰度值，从而能够加快收敛速度。而且，后面所介绍的几种技术能够有效地抑制星状伪迹，降低误差引入。

- **多级获取策略 MAS**

现有迭代算法采用的是逐方向更新策略，依次顺序对相邻投影方向的数据进行更新。因为顺序投影方向之间存在很大的相关性，容易引入误差。在安排投影数据的访问方式时应降低相邻投影数据之间的相关性，使各个相邻访问的投影方向尽可能地正交，从而加速收敛过程。为了满足这一要求，我们采用多级访问策略（MAS）重新对投影方向进行排序，使一系列被连续访问的投影角度在视角范围内尽可能地均匀分布。已经证明当两个方向夹角为 90°时它们的相关性最小，而第三个方向平分前两个方向夹角时三个方向相关性最小。基于这一点，MAS 可以对任意个数的投影方向进行重新排序。特别是当投影方向的个数是 2 的幂时，投影方向的顺序类似于一维快速傅里叶变换。假设投影方向个数为 P，每个方向依次编号为 0，1，……，P-1。MAS 将 P 个投影方向分成 L 级，L 可以表示为：$L=[\log_2 P]+1$。

在第一级，有两个投影方向分别为 0°和 90°；接下来的第二级也包括两个投影方向，分别为 45°和 135°；第三级有四个投影方向，依次为 22.5°、112.5°、67.5°和 157.5°。在 MAS 策略中，每一级的投影角度依次平分前面所有的投影角度。在电子断层扫描中，由于电镜本身的限制，投影方向无法覆盖 180°。我们将 MAS 策略进行改进，按照每个方向的编号进行重新排序。在第一级，选择编号为 0 和 P/2 的投影方向；第二级包括编号为 P/4 和（3/4）P 的两个投影方向；第三级包括编号依次为 P/8、（5/8）P、（3/8）P 和（7/8）P 的四个投影方向，后面的各级以此类推。在最后一级时，按照上述原则计算所得的方向可能已经被选择了，我们需要从该编号位置开始向两边找到距离最近且未被挑选的方向。在迭代开始之前，所有的投影角度按照 MAS 策略进行重排，然后每次迭代时按照新的投影角度顺序进行更新。这种策略能够大大降低相邻方向之间的相关性，从而提高算法的收敛速度。

- **自适应误差调整 ACC**

在 ASART 中，我们采用一种数据驱动的方式对 x_j 进行调整：第 k 次迭代投影误差反投影补偿时不仅考虑加权因子，而且还考虑 $x_j^{(k)}$自身的值，所以第 i 条射线投影对像素点灰度值 x_j 的误差补偿为：

$$\frac{w_{ij}x_j^{(k)}(p_i-\sum_{h=1}^{N}w_{ih}x_h^{(k)})}{\sum_{h=1}^{N}w_{ih}x_j^{(k)}}\ 。$$

- **列集合替代技术 CSS**

在 ASART 中，我们采用列集合替代技术 CSS，考虑使用一个标量 $\boldsymbol{\beta}$ 表示加权矩阵 $\boldsymbol{W}$ 每列和的最大值的倒数。而矩阵 $\boldsymbol{W}$ 中每列和的最大值可以认为是通过一个像素点射线的最大个数。而且根据 $\boldsymbol{W}$ 的计算模型，通过一个像素点的射线的最大个数即为投影方向的个数。由此，此标量可以表示为：$\boldsymbol{\beta}=1/\boldsymbol{B}$，其中，$\boldsymbol{B}$ 表示投影方向的个数。采用 CSS 技术，迭代过程不必再对 $\boldsymbol{W}$ 每一列求和，可以大大节约算法运行时间，同时也能节省存储空间。

根据以上介绍的四种关键技术，改进后的 ASART 算法迭代公式如下所示：

$$
\begin{cases}
x_j^{(0)} = \sum_{i=1}^{M} \dfrac{w_{ij} p_i}{\sum_{i=1}^{M} w_{ij}}, j = 1,2,\ldots,N \\
x_j^{(k+1)} = x_j^{(k)} + \lambda\beta \sum_{s=1}^{S} \dfrac{w_{ij} x_j^{(k)} \left(p_i - \sum_{h=1}^{N} w_{ih} x_h^{(k)}\right)}{\sum_{h=1}^{N} w_{ih} x_j^{(k)}}
\end{cases}
$$

我们已经分别在曙光 4000H、single GPU 和 multi-GPU 多种硬件平台上实现了 ASART 算法的并行化[16,21,22]。

3.3 基于 GPU 平台的电子断层重构软件 ATOM

针对现有电子断层成像技术中重构软件的不足，特别是迭代重构算法速度慢的缺点，我们开发了一款基于图形处理器（Graphic Processor Unit，GPU）平台的电子断层重构软件—ATOM[23]，实现了二维投影图像对位、确定重构参数、三维重构及二维数据的可视化等功能。其特点在于：在二维对位方面采用了迭代的平移旋转对位方法，提高对位精度；在三维重构方面，实现了 GPU 平台上的并行迭代重构算法，其中 Simultaneous Iterative Reconstruction Techniques (SIRT)算法达到 47 倍加速比。为保证与其他电镜软件的融合性，我们对软件操作过程中的电镜序列图像采用主流的 MRC 格式存储。ATOM 软件的设计流程如图 3 所示，其中实线框内为 ATOM 1.0 已实现的功能。

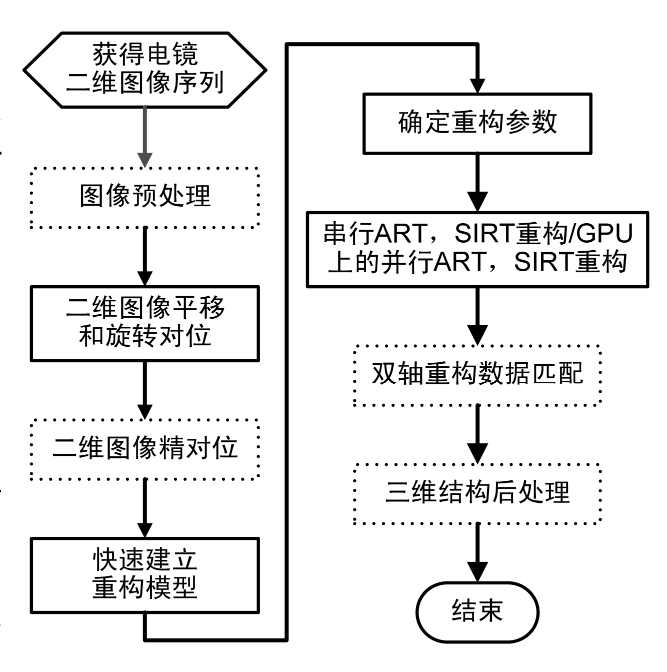


图 3. ATOM 软件基本流程

ATOM 1.0 是基于 Qt 和 CUDA 库开发的跨平台软件，使用 C++和 C 语言开发完成，其中界面代码由 C++完成，核心代码主要由 C 完成，通过编译可以运行在支持 Qt 和 CUDA 的所有操作系统上。ATOM 1.0 设计时将界面和可执行程序分离，用户可启动 ATOM 图形界面操作，也可直接执行可执行文件，在命令行下能完成所需功能。考虑到软件的易用和可理解性，我们将 ATOM 1.0 的图形界面设计为“基本功能”、“参



图 4. ATOM 界面

数交互"、"运行信息反馈"，以及"显示"4 个部分（如图 4）。ATOM 为开源软件，用户可以根据自己的实际需求，添加相应的代码，扩充相应功能。ATOM 的下载地址:
http://feilab.ibp.ac.cn/software/atom/atom.html.

目前 ATOM 已完成 2.0 版本的更新，图中的虚线框部分的功能也已实现。

4　总结与未来工作

本文详细介绍了电子断层三维重构技术的发展和现状，分析了当前存在的主要问题，着重介绍了我们在电子断层三维重构方面的相关研究工作：提出一种新的基于特征提取的二维投影图像对位算法 SAMA；实现了一种自适应迭代重构算法 ASART，能有效地加快收敛速度，得到更为精确的重构结果；开发了国内首款基于 GPU 平台的电子断层重构软件 ATOM。

在下一步研究工作中，我们将主要在以下四个方面继续深入开展研究：(1). 研究新的无胶体金二维图像对位算法，进一步提高对位参数的精度；(2). 在 ASART 基础上，继续进一步提高重构算法的收敛速度和精度；(3). 开发 ATOM2.0，研究在 GPU 集群上的电子断层三维重构并行；(4). 同生物学家合作，将电子断层三维重构技术与其他生物成像技术（如光学 PLAM 成像）相结合，开展光电联合三维成像的研究。

参考文献：

[1] 张凯, 张艳, 胡仲军等. 2010. 电子显微三维重构技术发展与前沿. *生物物理学报*, 26(7)：533-559.

[2] Sali A, Glaeser R, Earnest T, Baumeister W. 2003. From words to literature in structural proteomics. *Nature*, 422: 216-225.

[3] Zhang X, Jin L, Fang Q, et. al. 2010. 3.3Å Cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell*, 141: 472-482.

[4] 樊莉亚, 邵书伟, 王功明等. 2010. 单颗粒冷冻电镜三维重构研究的新进展. *信息技术快报*, 8(1)：1-10.

[5] Breakthrough of the year. 2002. *Sciences*, 298: 2297-2299.

[6] Frank J. 2006. Electron tomography: methods for three-dimensional visualization of structures in the cell. *Second edition, New York: Springer*. 135-139.

[7] Zheng SQ, Keszthelyi B, Branlund E, et. al. 2007. UCSF tomography: An integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. *Journal of Structural Biology*, 157(1): 138-147.

[8] Kremer JR, Mastronarde DN, McIntosh JR. 1996. Computer visualization of three-dimensional image data using IMOD. Journal of Structural Biology, 116: 71-76.

[9] Winkler H. 2007. 3D reconstruction and processing of volumetric data in cryo-electron tomography. *Journal of Structural Biology*, 157(1): 126-137.

[10] Frank J, Shimkin B, Dowse H. 1981, Spider-a modular software system for electron image processing. *Ultramicroscopy*, 6(4): 343-357.

[11] Andersen AH, KaK AC. 1984. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6: 81-94.

[12] Sorzano C., Marabini R., Boisset N., et. al. 2001. The effect of overabundant projection directions on 3D reconstruction algorithms. *Journal of Structural Biology* 133(2-3): 108-118.

[13] 单桂华, 刘俊, 田东等. 2010. 冷冻电镜密度图可视化软件 VAT4M. *生物物理学报*, 26(7)：590-599.

[14] Fernandez JJ. 2008. High performance computing in structural determination by electron cryomicroscopy. *Journal of Structural Biology*, 164: 1-6.

[15] Agulleiro JI., Fernandez JJ. 2011. Fast tomographic reconstruction on multicore computers. *Bioinformatics*, 27: 582-583.

[16] Wan XH., Zhang F., Liu ZY. 2009. Modified simultaneous algebraic reconstruction technique and its

parallelization in cryo-electron tomography. *ICPADS' 2009*.

[17] Vazquez F., Garzon EM., Fernandez JJ., 2010. A matrix approach to tomographic reconstruction and its implementation on GPUs. *Journal of Structural Biology*, 170: 146-151.

[18] Xu W., Xu F., Jones M., et. al. 2010. High-performance iterative electron tomography reconstruction with long-object compensation using graphics processing units (GPUs). *Journal of Structural Biology*, 171: 142-153.

[19] Chu Q.，Zhang F.，Zhang K.，et. al. An accurate, automatic method for markerless alignment of electron tomographic images. *BIBM2010*.

[20] Wan XH., Zhang F. Chu Q., et al, 2011. Three-dimensional Reconstruction by Adaptive Simultaneous Algebraic Reconstruction Technique in Electron Tomography. *Journal of Structural Biology*, 175: 277-287.

[21] Wan XH., Zhang F., Chu Q. and Liu ZY. 2011. High-performance Blob-based Iterative Reconstruction of Electron Tomography on Multi-GPUs. ISBRA2011, LNCS 6674:61-72.

[22] Wan XH., Zhang F., Chu Q. and Liu ZY. 2011. High-performance Blob-based Iterative Three-dimensional Reconstruction of Electron Tomography using Multi-GPUs. Submitted to BMC Bioinformatics.

[23] 储琪, 张凯,万晓华等. 2011. ATOM 1.0：基于 GPU 的电子断层重构软件. *生物物理学报*, 27(3)：231-241.

作者简介：

张　法： 中国科学院计算技术研究所前瞻研究实验室，副研究员，zf@ncic.ac.cn

万晓华： 中国科学院计算技术研究所前瞻研究实验室，研究生

储　琪： 中国科学院计算技术研究所前瞻研究实验室，研究生